

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
1. April 2004 (01.04.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/027707 A2

(51) Internationale Patentklassifikation⁷: **G06T 1/00**

(21) Internationales Aktenzeichen: PCT/EP2003/010030

(22) Internationales Anmeldedatum:
10. September 2003 (10.09.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 42 087.4 11. September 2002 (11.09.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme
von US): DAIMLERCHRYSLER AG [DE/DE]; Ep-
plestr. 225, 70567 Stuttgart (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): UWE, Franke

[DE/DE]; Benzstrasse 33, 73066 UHINGEN (DE). PAYSAN,
Pascal [DE/DE]; Katharinenstrasse 35, 73066 UHINGEN
(DE). STEIN, Fridtjof [DE/DE]; Helene-Lange-Strasse
52, 73760 Ostfildern (DE). WÜRZ-WESSEL, Alexander
[DE/DE]; Ziegelstrasse 66, 73084 Salach (DE).

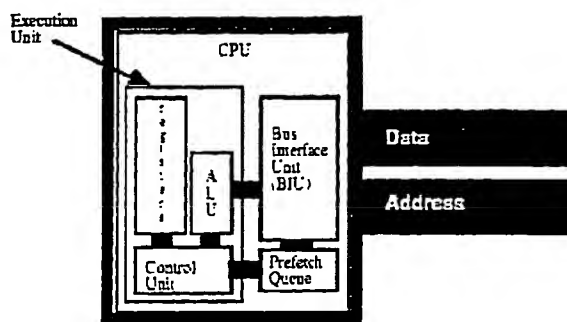
(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,
CU, CZ, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG,
SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,

[Fortsetzung auf der nächsten Seite]

(54) Title: IMAGE PROCESSING ON HARDWARE THAT IS OPTIMIZED WITH REGARD TO VECTOR CALCULATION
AND COLOR MIXTURE

(54) Bezeichnung: BILDVERARBEITUNG AUF FÜR VEKTORRECHNUNG UND FARBMISCHUNG OPTIMISierter HARD-
WARE



(57) Abstract: Algorithms for processing digitized images (digitally recorded, for example, by digital CCD or CMOS cameras, or analogally recorded by analog CCD cameras and subsequently converted in analog/digital converters) and for processing the information contained therein require, according to use, an extremely high amount of computing. Computer hardware, which is specialized for vector calculation and color mixture, so-called graphical processing units GPU (e.g. so-called 3D-accelerated graphics cards, preferably having the OpenGL standard and/or Direct X standard that are/is implemented on the hardware), and which has been developed particularly for entertainment electronics (computer games and animation) or for the depiction of three-dimensional structures (CAD), includes realizations of image processing algorithmics based on semiconductor technology. The inventive implementation of an item of hardware of the aforementioned type as an essential part of image processing enables an economically sensible realization of a computer performance required for carrying out real-time image processing. According to the invention, the GPU (also denoted, in part, as a VPU video processing unit) is not used in a conventional manner (namely for transforming artificial three-dimensional worlds into a two-dimensional view that is as realistic as possible on a display screen, but instead is used for the contrary by transforming real two-dimensional views (generally speaking: camera images) in order to determine the actual geometries prevailing in the real three-dimensional space.

[Fortsetzung auf der nächsten Seite]

WO 2004/027707 A2



TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

Veröffentlicht:

- *ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts*

(57) Zusammenfassung: Algorithmen zur Bearbeitung von digitalisierten Bildern (digital aufgenommen beispielsweise durch digitale CCD- oder CMOS-Kameras, oder analog aufgenommen durch analoge CCD-Kameras und anschließend in Analog-Digital-Wandler konvertiert) und der darin enthaltenen Informationen sind je nach Anwendung extrem rechenaufwendig. Auf Vektorrechnung und Farbmischung spezialisierte Computer Hardware sogenannte Graphical Processing Units GPU (z.B. sog. 3D-beschleunigte Grafikkarten, vorzugsweise mit hardwareseitig implementiertem OpenGL- und / oder Direct-X-Standard), die vornehmlich für die Unterhaltungselektronik (Computerspiele und -animation) oder die Darstellung von 3-dimensionalen Konstruktionen (CAD) entwickelt wurde, beinhalten Realisierungen von Bildverarbeitungsalgorithmik auf Basis der Halbleitertechnik. Durch die erfinderische Implementierung einer solcher Hardware als wesentlicher Teil der Bildverarbeitung, kann eine für eine Echtzeit-Bildverarbeitung notwendige Rechnerleistung wirtschaftlich sinnvoll realisiert werden. In erfinderischer Weise wird hierbei die GPU (teilweise auch als VPU - Video Processing Unit bezeichnet) nicht in bisher üblicher Weise verwandt (nämlich zur Transformation von synthetischen 3-dimensionalen Welten in eine möglichst realistische 2-D Ansicht auf einem Bildschirm), sondern gerade umgekehrt, nämlich zur Transformation realer 2-D Ansichten (im allgemeinen: Kamerabilder) um daraus die tatsächlich im realen 3-dimensionalen Raum vorherrschenden Geometrien zu ermitteln.

5 Bildverarbeitung auf für Vektorrechnung und Farbmischung op-
timierter Hardware

Die Erfindung betrifft eine Vorrichtung und ein Verfahren
10 nach dem Oberbegriff der Patentansprüche 1 und 5.

Algorithmen zur Bearbeitung von digitalisierten Bilder (digi-
tal aufgenommen beispielsweise durch digitale CCD- oder CMOS-
Kameras, oder analog aufgenommen durch analoge CCD-Kameras
15 und anschließend in Analog-Digital-Wandlern konvertiert) und
der darin enthaltenen Informationen sind je nach Anwendung
extrem rechenaufwendig.

Bisher wird die Bildverarbeitung im allgemeinen auf frei pro-
20 grammierbaren Recheneinheit (Central Processing Unit, CPU)
implementiert. Die Rechenkapazität der CPU ist hierbei der
limitierende Faktor für die Informationsaufbereitung und -ge-
winnung. Die Anwendung redundanter oder ergänzender Algorith-
men etwa zur wechselseitigen Kontrolle der Ergebnisse ist aus
25 zeitlichen Gründen oft nicht möglich. Viele Applikationen un-
terliegen der „Video-Echtzeitanforderung“ , d.h. Abarbeitung
der Informationen eines Bildes in der Zeit, bis das nächste
Bild zur Verfügung steht. Viele Kameras arbeiten mit einer
Bildrate von 25 Hz, d.h. es stehen 40ms an Rechenzeit zur
30 Verfügung. Da durch die verwendete CPU auch die Rechenlei-
stung begrenzt ist, sind der Informationsauswertung klare
Grenzen gesetzt, selbst im Hinblick darauf, dass in den letz-
ten Jahren Prozessoren mit stetig steigende Rechenleistung
auf den Markt gekommen sind. Aus diesem Grund sind viele Be-
35 arbeitungsschritte, die zwischenzeitlich algorithmisch be-

herrscht werden nicht einsetzbar, da der Zeitaufwand zu hoch ist. Insbesondere die simultane Anwendung aufwendiger Algorithmen, z.B. zur Bildverbesserung (Farbwertinterpolation), Informationsaufbereitung (z.B. Umrechnung einer Ansicht in
5 eine sog. Vogelperspektive), Korrektur von Bildfehlern (z.B. Verzeichnung durch Linsen- oder Einbaufehler) und zur Informationsgewinnung (Auswertungsalgorithmik, wie Fahrspurerkennung oder Objektklassifikation) sind bei aktuellen Rechnerleistungen wirtschaftlich nicht sinnvoll realisierbar.

10

Um eine ausreichende Verarbeitungsgeschwindigkeit zu gewährleisten ist es deshalb, insbesondere aus der Militärtechnik, auch bekannt, spezielle auf die vorliegende Bildverarbeitungsaufgabe spezifizierte Prozessoren (ASICs, application
15 specific integrated circuits) zu verwenden. Bei solchen Prozessoreinheiten ist es möglich eine spezifische Verarbeitungsalgorithmik hardware-technisch zu realisieren und auf diese Weise eine hohe Verarbeitungsgeschwindigkeit zu erzielen. Eine solche Lösung birgt jedoch den Nachteil hoher Gestehungskosten, da nur eine anwenderspezifische Kleinserie gefertigt
20 werden kann, und einer starren Verarbeitungsstruktur, welche sich nur sehr eingeschränkt an sich ändernde Anforderungen, welche sich beispielsweise aus den im Rahmen des Produktentwicklungsprozess entstehenden Erkenntnissen ergeben, anpassen
25 lassen.

Aufgabe der Erfindung ist es deshalb, eine Vorrichtung und ein Verfahren zur Bildverarbeitung zu finden, welche wirtschaftlich sinnvoll, hohe Rechnerleistung und flexible Produktgestaltung garantiert.
30

Die Aufgabe wird durch eine Vorrichtung und ein Verfahren mit den Merkmalen der Patentansprüche 1 und 5 gelöst.

Vorteilhafte Ausgestaltungen und Weiterbildungen der Erfindung werden durch die abhängigen Patentansprüche beschrieben.

Auf Vektorrechnung und Farbmischung spezialisierte Computer-Hardware sogenannte Graphical Processing Units GPU (z.B. sog. 3D-beschleunigte Grafikkarten, vorzugsweise mit hardwareseitig implementiertem OpenGL- und / oder Direct-X-Standard), die vornehmlich für die Unterhaltungselektronik (Computerspiele und -animation) oder die Darstellung von 3-dimensionalen Konstruktionen (CAD) entwickelt wurde, beinhalten Realisierungen von Bildverarbeitungsalgorithmik auf Basis der Halbleitertechnik. Durch die erfinderische Implementierung einer solcher Hardware als wesentlicher Teil der Bildverarbeitung, kann eine für eine Echtzeit-Bildverarbeitung notwendige Rechnerleistung wirtschaftlich sinnvoll realisiert werden. In erfinderischer Weise wird hierbei die GPU (teilweise auch als VPU - Video Processing Unit bezeichnet) nicht in bisher üblicher Weise verwandt (nämlich zur Transformation von synthetischen 3-dimensionalen Welten in eine möglichst realistische 2-D Ansicht auf einem Bildschirm), sondern gerade umgekehrt, nämlich zur Transformation realer 2-D Ansichten (im allgemeinen: Kamerabilder) um daraus die tatsächlich im realen 3-dimensionalen Raum vorherrschenden Geometrien zu ermitteln.

Die Erfindung macht somit zwar Gebrauch von einer allgemein bekannten Vorrichtung, nämlich einer GPU, nutzt diese aber in absoluter Abkehr von deren bestimmungsgemäßen Gebrauch. Insbesondere die Verbindung des Ausgangs einer Kameraeinheit mit dem Eingang einer GPU bzw. einer Graphikkarte ist im Rahmen dieser Erfindung erstmalig und wird dem Fachmann durch den Stand der Technik auch nicht nahegelegt. Im Gegenteil, zeigt der Stand der Technik zur Optimierung der Verarbeitungskapazität einer Kamerabildverarbeitung die Möglichkeit

auf, die Kamera mit einer intelligenten Frame-Grabber-Karte zu kombinieren, um die Bilddaten vor der eigentlichen Bildverarbeitung durch die CPU des Rechners möglichst effektiv vorzuverarbeiten. Solche Frame-Grabber-Karten (beispielsweise Snapper-PCI-8/24 der Fa. Active Silicon) ermöglichen es, beispielsweise mittels eines sogenannten Data-Mappers gezielt Ausschnitte aus den Bilddaten auszuwählen, nachfolgend der Bildverarbeitung zugeführt werden sollen; eine Verarbeitung der Bilddaten durch eine GPU findet hierbei jedoch nicht statt.

Durch recht freie Programmierbarkeit der GPU wird es auch möglich, ein solch erfinderisches Bildverarbeitungssystem flexibel an sich ändernde Bedingungen anzupassen. Auch können durch die Erfindung bildverarbeitende Geräte zuverlässiger und ausfallsicherer gemacht werden, da beispielsweise die allgemein übliche Parallel-Schaltung mehrerer CPU's zur Leistungssteigerung entfällt. Andererseits erlaubt es die erhöhte Rechenleistung, mehr Information in gleicher Zeit zu verarbeiten, beziehungsweise innerhalb der zur Verfügung stehenden Zeit bestimmte Informationen zu Verifizieren, indem die Bilddaten parallel durch mehrer Algorithmen verarbeitet werden, welchen unterschiedliche Lösungsansätze zu Grunde liegen. Die Verwendung einer GPU für Bildverarbeitungsaufgaben bildet auch einen wirtschaftlich optimalen Kompromiss zwischen der herkömmlichen Verwendung von CPU's und der bekannten spezifischen Hardwarerealisierung der Algorithmik auf einem AISIC, da so für andere Zwecke (meist Computerspiele oder CAD-Anwendungen) konzipierte Standard-Hardware, welche mittlerweile als kostengünstige Massenware zur Verfügung steht, verwendet werden kann.

Die wesentlichen Unterschiede zwischen einer CPU wie sie in allen gängigen Computern als zentrale Rechneinheit zu fin-

den ist und einer GPU, wie sie insbesondere auf Graphikkarten zur Wandlung synthetischer 3-dimensionaler Bilddaten in 2-dimensionale Darstellungen Verwendung findet, sei zur weiteren Klarstellung der Erfindung nachfolgend auch anhand der
5 Figuren 1 und 2 erläutert.

Eine CPU, wie sie in Figur 1 aufgezeigt ist, besteht aus einer Kontroll-Einheit zu deren Steuerung, einer arithmetischen-logischen Einheit (ALU), Registern (die dauerhaft existieren und mit beliebigen Daten gefüllt werden können) ,
10 einer Prefetch-Queue, die Daten vorausschauend über den Daten-Bus in den Cache holt und dem Daten-Bus-Interface, welches die Schnittstelle zur restlichen Hardware (Grafikkarte, Schnittstellenkarte, Festplatte, CD-ROM, etc.) darstellt. Die
15 CPU ist auf die sequentielle Bearbeitung von Daten ausgelegt.

Figur 2 zeigt deutlich, dass die einzelnen Einheiten der GPU in Form einer Pipeline angeordnet sind. Die einzelnen Daten werden hierdurch sequentiell durch die Einheiten, den sogenannten Shadern, der GPU geschoben. Die wesentlichen Elemente einer GPU lassen sich wie folgt beschreiben:

- Die Vektorrechnung wird in der sogenannten Geometrie-Pipeline vorgenommen. Diese sind in der Lage dreidimensionale Vektoren zu verarbeiten. Sollte die standardmäßig implementierten Methoden für den jeweiligen Algorithmus nicht
25 ausreichen, können diese durch den Programmierer durch spezifische Programme (Vertex-Shader) erweitert und angepasst werden. Dadurch können jedem Pixel bestimmte Projektionsentfernungen und Rechnung zugeordnet werden und dadurch
30 Entzerrungen und Perspektiven rekonstruiert werden. Jede Art von Bildverarbeitungsalgorithmik, die Vektorrechnung beinhaltet (z.B. dreidimensionale Rekonstruktion einer Szene durch Auswertung von Stereobildern) kann von diesen Ein-

heiten unterstützt und / oder ersetzt werden. Konkrete Anwendungen siehe unten.

- Die Farbmischung findet auf sogenannten Textur- und Pixel-Shadern statt. Diese Shader können Farbwerte nach bestimmten Muster innerhalb des Bildes verschieben oder farbliche Zwischenwerte für bestimmte Bereiche generieren. Es stehen bei den derzeit auf dem Markt verfügbaren GPU bis zu 4 Textur-Einheiten zur Verfügung, so dass z.B. auf eine Geometrie mehrere Texturoperationen gleichzeitig angewendet werden können. Zudem können mehrere Einheiten arithmetisch miteinander verbunden werden. Operationen die etwa auf dem Vergleich von Farbwerte oder dem Positionsvergleich identischer Farbwerte oder Farbwertkombinationen in verschiedenen Bildern beruhen können von diesen Einheiten unterstützt und / oder ersetzt werden.

Innerhalb der Pipeline können Elemente verworfen werden. Dies geschieht im Prinzip für jedes Element einzeln, d.h. ohne Kenntnis benachbarter Elemente. Dies hat folgende Konsequenzen:

- Da auch keine dauerhaften Register wie bei der CPU zur Verfügung stehen, ist z.B. keine Aufsummieren von Farbwerten über einen Bildbereich möglich. Register werden nur während der Berechnung eines Elements (z.B. eines Vektors) angelegt und danach wieder gelöscht (dies könnte sich evtl. in den nächsten Generationen von GPU's ändern.)
- Da die einzelnen Elemente unabhängig voneinander bearbeitet werden, ist die Reihenfolge der Bearbeitung irrelevant. Dies ist ein wesentliche Voraussetzung für die Parallelisierung der Bearbeitung und die Geschwindigkeiten, die damit erreicht werden können. Auf der Grafikkarte können deshalb mehrere Einheiten untergebracht werden. Gängige GPU's bzw. Graphikkarten haben beispielsweise bis zu 4 Textur-

Einheiten (zukünftig werden vermutlich noch mehr Einheiten vorhanden sein), die parallel arbeiten.

- Einheiten wie die ALU oder die Prefetch-Queue sind im Gegensatz zur CPU bei einer GPU nicht vorhanden.

5

Die oben angeführten Unterschiede zwischen der GPU und der im Stand der Technik für die Bildverarbeitung herangezogene CPU erfordert eine spezielle Anpassung der bisher bekannten Bildverarbeitungsalgorithmik an die vorliegende GPU-Architektur.

- 10 Um den Begriff GPU zusätzlich zu den vorigen Erläuterungen eindeutig zu definieren, seien nachfolgend beispielhaft gängige für die erfindungsgemäße Verwendung zur Bildverarbeitung geeignete GPU's angeführt:

- GeForce4 Serie der Firma NVIDIA Corporation
- 15 - nForce Serie der Firma NVIDIA Corporation
- Parhelia-512 GPU der Firma Matrox Graphics Inc..
- ATI RADEON GPU der Firma ATI Technologies Inc..
- Wildcat VP Visual Processing Unit (VPU) der Firma 3D-Labs Inc.

20

Selbstverständlich wird, insbesondere auch im Rahmen dieser Anmeldung unter einer GPU nicht nur der isolierte Prozessorbaustein als solcher (Beispiele siehe oben), sondern auch die mit diesem Graphik Prozessor zusammenhängende Hardware-

- 25 Peripherie (beispielsweise, diesem zugeordnete Bildspeicher oder auch speziell insbesondere für die Systemverwaltung optimierte CPUs) verstanden. Solche Systeme aus GPU-Prozessor mit zugehöriger Hardware-Peripherie wird gemeinhin auch als Graphikkarte bezeichnet, da es meist als ein in ein CPU-
- 30 basiertes Computersystem integrierbare Einsteckkarte realisiert wird. An Stelle solcher Graphikkarten ist es aber gleichwohl auch möglich, dass solche Systeme aus GPU-Prozessor und Hardware-Peripherie als eine gemeinsame Einheit

direkt auf dem Motherboard des Computersystems realisiert sind.

In besonders vorteilhafter Weise erhält das erfinderischen
5 Bildverarbeitungssystem die zu verarbeitenden Bilddaten direkt durch eine Kameraeinheit geliefert, wodurch sich ein echtzeitfähige Kamerabildverarbeitung realisieren lässt, insbesondere zum Einsatz in einem Kraftfahrzeug im Zusammenhang mit einem Sicherheits- oder Assistenzsystem. Andererseits ist
10 es sehr wohl auch möglich das Bildverarbeitungssystem zur effektiven und schnellen Nachbearbeitung von in einem Bildspeicher gespeicherten Kamerabildern heranzuziehen.

Je nach Anwendungsfall können sodann die Auswertungsergebnisse der Bildverarbeitung sowohl auf einer Bildanzeige beispielsweise einem Kraftfahrzeugführer dargebracht werden, oder auch in besonders vorteilhafter Weise an andere Systemeinheiten weitergeleitet werden, welche auf dieser Grundlage auf andere Systeme einwirken, um beispielsweise ein Fahrzeug
20 vor einem aus den Bilddaten erkannten Hindernis automatisch abzubremsen.

Zur weiteren Leistungssteigerung des erfinderischen Bildverarbeitungssystems ist es gewinnbringend, wenn parallel zu dem
25 Graphikprozessor eine frei programmierbare CPU angeordnet ist. Eine solche CPU ist insbesondere für organisatorische Aufgaben im Rahmen der Bildverarbeitung geeignet und könnte in steuernder oder regelnder Weise auf den Verarbeitungsablauf in der GPU einwirken. So könnten beispielsweise ausgehend von den Auswertungsergebnissen oder von veränderten Rahmenbedingungen (beispielsweise: Wechsel zu einer anderen Kameraoptik) bestimmte Parameter der im wesentlichen auf der GPU implementierten Verarbeitungsalgorithmik verändert werden. Andererseits lässt sich durch eine solche Parallelstel-

lung einer CPU zu der eigentlichen graphischen Verarbeitungseinheit GPU die Verarbeitungseffizienz steigern, indem anstelle der gemeinhin üblichen Abfolge der Bildverarbeitung, dergestalt dass Bild per DMA von der Kamera direkt in die GPU und nach der Aufbereitung durch die Recheneinheiten der GPU über DMA in den Hauptspeicher der CPU zur Weiterverarbeitung geladen werden, ganz oder teilweise umgekehrt werden. So ist es in vorteilhafter Weise möglich, wie bisher, zuerst die Bilddaten von der Kamera über DMA in den CPU-Hauptspeicher zu laden. Hierzu werden nach ersten Verarbeitungsschritten durch die CPU die Bilddaten und/oder Teile davon in einen oder mehrere der GPU zugeordnete Bilddaten-Speicher verschoben, verarbeitet und gegebenenfalls auch angezeigt. Nach der Bearbeitung durch die GPU, können die resultierenden Daten wieder in den Hauptspeicher der CPU geladen werden. Diese wechselseitige Bearbeitung durch CPU und GPU kann je nach Anforderungen des Algorithmus oder nach Kriterien einer effizienten Implementierung mehrfach geschehen.

Wie bereits zuvor erwähnt, müssen aus dem Grund, dass eine GPU im Gegensatz zu einer CPU nicht frei programmierbar ist, die eventuell bereits aus dem Stand der Technik bekannten Teilalgorithmen zur Bildverarbeitung abgeändert und auf eine GPU-spezifische Implementierung angepasst werden. In besonders vorteilhafter Weise können hierdurch die speziell auf die Shader-Strukturen der GPU angepassten Algorithmen auch die durch die Vertex- oder Textur-/Pixel-Shader zur Verfügung gestellte Funktionalität ausnutzen.

Nachfolgend wird an mehreren Beispielen das notwendige erfinderische Vorgehen zur Anpassung von Bildverarbeitungsalgorithmen auf eine GPU-Hardware-Struktur beschrieben.

Entzerrung von Bildern

Durch Linsenfehler und mangelhaften Einbau von Linsen und Kamerachips entstehen Verzeichnungen (Abbildungsfehler) die mittels der GPU, insbesondere deren Pixel-Shadern, unter Verwendung von Kalibrationsalgorithmen schnell und effektiv korrigiert werden können. Hierbei werden aus den Bilddaten der Kamera die Verzeichnungsparameter bestimmt und hieraus die notwendigen Rechenvorschriften bestimmt, wie Farb- oder Grauwert in den Bilder verschoben werden müssen, für die Kalibrationsalgorithmen, damit die Bilddaten nach Durchlauf dieser Stufe der Bildverarbeitung so vorliegen, als ob sie von einer idealen Kamera entstammten.

Farbmischung

Die Farbmischung findet auf sogenannten Textur- und Pixel-Shadern statt. Diese Shader können Farbwerte nach bestimmten Muster innerhalb des Bildes verschieben oder farbliche Zwischenwerte für bestimmte Bereiche generieren. Es stehen bis zu 4 Textur-Einheiten zur Verfügung, so dass z.B. auf eine Geometrie (siehe oben) mehrere Texturoperationen angewendet werden können. Zudem können mehrere Einheiten arithmetisch miteinander verbunden werden. Operationen Bildverarbeitung im weitesten Sinne, die etwa auf dem Vergleich von Farbwerte oder dem Positionsvergleich identischer Farbwerte oder Farbwertkombinationen in verschiedenen Bildern beruhen können von diesen Einheiten unterstützt und / oder ersetzt werden.

Interpolationsverfahren

Die Ermittlung von Farb- oder Grauzwischenwerten, etwa bei einer Bildgrößenänderung oder im Rahmen der Entzerrung von Bilddaten, sind sehr aufwendig und können bisher nicht oder nur partiell (auf einzelnen, vorher bestimmten Bildbereichen) durchgeführt werden. Die fest implementierten Interpolationsfunktionen, insbesondere die Funktionalität der Pixel-Shader, verbessern die gesamten Bilddaten, so dass verschiedene Algo-

rithmen (etwa Kantenextraktionen), die auf entsprechend feine Strukturen angewiesen sind, bessere Ergebnisse liefern.

Um Farbwerte benachbarter Pixel aufsummieren zu können, kann
5 die Interpolationsroutine entgegen ihrer eigentlichen Bestimmung in erfinderischer Weise so genutzt, dass zur gewünschten Interpolation über jeweils 4 Pixel der Bilddaten eines Bildausschnittes dieser Bildausschnitt in beiden Dimensionen um den Faktor 2 verkleinert wird, wobei die GPU standardmäßig
10 eine Interpolation der Bilddaten über die 4 jeweils benachbarten Pixel ausführt. Da dieser Faktor 4 jedoch bekannt ist, lässt sich auf einfache Multiplikation der aus der Interpolation resultierenden Farbwerte der Bilddaten die eigentlich gesuchte Farbwertsumme wieder herstellen. Dieses Vorgehen
15 lässt sich selbstverständlich entsprechend auch mit GPU-Hardware anwenden, welche standardmäßig über mehr als 4 umliegende Pixelwerte interpoliert.

Rekonstruktion von Bildern

20 Spezielle Kamerageometrien, wie z.B. panoramische Kameras, bestehend aus Spiegeln und Linsen, erfordern geometrische Modelle zur Rekonstruktion des Bildes. Wird z.B. eine Rotationsparaboloid in der panoramischen Kamera verwendet, so lässt sich ein Vektormodell des Rotationsparaboloids in der GPU ab-
25 legen. Das direkt über DMA in die GPU geholte Bild wird auf diese Vektorfläche gelegt; dies geschieht insbesondere im Vertex-Shader. Da die äußere Geometrie der Vektorfläche der inneren Geometrie des aufgenommenen Bildes entspricht, findet automatisch Rekonstruktion des Bildes statt. Die aufgenommene
30 Bildgröße, die Blickrichtung und der Kamerapunkt innerhalb der geometrischen Anordnung können frei angeben. Dadurch lassen sich Bilddaten beliebiger Kamera-Ansichten innerhalb des aufgenommenen Raumes in diesen rücktransformieren.

Wird beispielsweise bei der Bildaufnahme durch eine panoramische Kamera als Spiegel nicht ein Paraboloid, sondern ein Hyperboloid, Ellipsoid, eine Kugel oder eine andere geometrische Form verwendet, so lassen sich diese auf die ein und dieselbe Art und Weise in der GPU hinterlegen und entsprechende Ergebnisse erzielen.

10 Rekonstruktion von Freiflächen-Reflexionen

Der Einsatz von Vektorflächen ist nicht auf die unter dem vorherigen Punkt „Rekonstruktion von Bildern“ beschriebenen wohldefinierten Geometrien beschränkt. Da im Prinzip für jedes Bildelement (Pixel) ein spezieller Vektorpunkt angegeben werden kann, können auch Spiegelungen an bekannten Freiflächen mittels der GPU rücktransformiert werden. So kann beispielsweise aus den Stereobilddaten der tatsächliche Standort eines sich auf der Motorhaube eines Kraftfahrzeuges spiegelnden Baumes durch Rücktransformation errechnet werden, da die Abbildungsgeometrien der Motorhaube, beispielsweise aus deren CAD-Daten, genau bekannt sind.

Sind die Vektordaten für die angegeben Fläche nicht dicht über den Pixeln (es wird z.B. zur Reduzierung der zu übertragenden Vektorgeometrie-Daten nur ein Vektorpunkt für jedes vierte Pixel angegeben), so kann die GPU in besonders vorteilhafter Weise genutzt werden, um die Zwischenwerte für die nicht angegebenen Pixel automatisch linear durch die bereits im Vertex-Shader realisierte Hardware zu interpolieren. Dies bedeutet, dass automatisch durch die GPU-Hardware zwischen den gegebenen Flächenpunkten eine Triangulation stattfindet. Aufwendige Ansätze, die zuerst in den Daten entsprechende Dreiecksberechnung (Triangulär-Mesh) durchführten, können somit entfallen.

Berechnung von anderen Kameraperspektiven

Für bestimmte Anwendungen kann es sinnvoll sein, den Blickpunkt der Kamera virtuell zu verlegen, beispielsweise um eine Ansicht in Vogelperspektive zu schaffen. Ein anderes Beispiel hierzu ist die Erkennung von Fahrbahn- und fahrbahnparallelen Strukturen (Spurerkennung) aus Kamerabildern. Üblicherweise wird die Kamera zur Spurerkennung im Frontbereich des Fahrzeugs eingebaut, so dass der Blick der Kamera in etwa dem Blick des Fahrers entspricht. Da die normale Linsenkamera perspektivisch arbeitet, werden die Fahrbahnstrukturen als auf den Fluchtpunkt zulaufenden Strukturen im Bild wiedergegeben. D.h. in der Realität parallele Strukturen sind im Bild nicht mehr parallel. Ist die Einbauhöhe, der Blick- und Nickwinkel der Kamera bekannt, so kann das Bild in eine sog. Vogelperspektive umgerechnet werden. Dazu werden die Bildinformationen algorithmisch auf die Straßenebene zurückprojiziert. Das so entstandene Bild wird durch eine virtuelle Kamera, die sich geometrisch vor dem Fahrzeug befindet und deren Bildebene parallel zur Straßenebene ist abgebildet. Durch dieses Verfahren werden parallele Strukturen in der Straßenebene bzw. in dazu parallelen Ebenen als parallele Strukturen im aus der Bildverarbeitung resultierenden Datensatz abgebildet. Da es sich bei dieser Vogelperspektiven-Berechnung um reine geometrische Modelle handelt, kann hierzu in gewinnbringender Weise die auf graphische Transformation optimierte Hardware der GPU, insbesondere der Vertex-Shader, genutzt werden. Für dieses Vorgehen insbesondere zwei Verfahrensabläufe denkbar:

- Beim einen gewinnbringenden Verfahrensablauf werden in einem ersten Schritt eine Ebene festgelegt, die der Straßenebene entspricht. Hierzu das Ergebnis einer Nickwinkelbestimmung der Kamera zur Definition der Ebene berücksichtigt werden, d.h. es ist eine ‚Nachführung‘ der Ebene während der Bewegung

der Kamera möglich, beispielsweise während der Fahrt einer Kamera in einem Kraftfahrzeug. Das Kamerabild wird nun ganz oder in Teilen auf die Ebene gelegt. Die von der GPU zur Darstellung automatisch durchgeführte Projektion erledigt die
5 Rekonstruktionsaufgabe.

- Zum anderen ist es in ebensolch vorteilhafter Weise möglich, dass auf einer Ebene der Projektion, d.h. dem Nick- und Blickwinkel, entsprechend die Textur-Koordinaten des Bildes verändert werden. Dies erzielt den gleichen Rekonstruktions-
10 effekt.

Weitere Anwendung solcher Rekonstruktionen neben der Spurerkennung sind etwa Rückfahrkameras für Pkw, Transporter oder Lkw. Dabei wird das aufgenommene Bild in die Perspektive ei-
15 ner Kamera umgerechnet, die hinter dem Fahrzeug zu schweben scheint. Eine weitere Anwendung wäre die Überwachung des toten Sichtbereichs direkt vor einem Lkw. Durch die hohe Sitzposition des Fahrers und die hohe Scheibenunterkante ist der Raum unmittelbar vor dem Fahrzeug für den Fahrer nicht ein-
20 sehbar. Durch eine oben angebrachte Kamera kann dieser Raum überwacht werden. Eine hilfreiche Darstellung des Bildes für den Fahrer ist wiederum eine Vogelperspektive, bei der die virtuelle Kamera vor dem Fahrzeug ,schwebt'. In den aufgezählten Beispielen wurde als Bildquelle eine normale, perspektivische Linsenkamera angeführt. Es können jedoch genauso
25 Bilddaten wie sie von einer panoramische und/oder orthographische Kameras geliefert werden, verwendet werden.

Korrelation von Bildern

30 Verfügt eine GPU über mehrere Textur-Shader, können auf einfache Weise schnelle Korrelationsroutinen, beispielsweise für die Stereobildverarbeitung auf der GPU realisiert werden. Durch geeignete Verschiebung der Arbeitspunkte, berechnet man an entsprechend versetzten Stellen in den jeweiligen Textur-

Shadern die Farbwerte. Nach dieser Berechnung werden die hierbei gewonnenen Texturen geeignet gemischt, um ein Korrelationsergebnis, beispielsweise ein Entfernungsbild bei der Stereobildverarbeitung zu erhalten.

5

Realisierung der Kamera-Belichtungsregelung bzw. Farbwert-
10 Anpassung durch die GPU.

Moderne CMOS- und CCD-Kameras haben einen großen Dynamikbereich, d.h. es können lichtschwache und lichtstarke Objekte gleichzeitig abgebildet werden. Um diesen großen Helligkeitsbereich abzudecken, verwenden die Kamera eine 12 Bit-
15 Repräsentation für das Bild. Da viele Displays nur 8 Bit-Grauwertbilder anzeigen können, müssen die Kamerainternen 12 Bit auf 8 Bit reduziert werden. Dies kann auf unterschiedliche Weise geschehen. Zum einen kann ein 8 Bit Bereich innerhalb der 12 Bit festgelegt werden. Dann wird entsprechend die
20 Belichtungszeit der Kamera geregelt, damit ein verwertbares Bild entsteht. Die Belichtungszeit wird durch Auswertung der Intensitätsverteilung im ganzen Bild oder bestimmten Bildteilen berechnet. Dies geschieht durch die CPU. Zum anderen kann der 8 Bit Bereich innerhalb der 12 Bit variabel gewählt werden. Dann besteht die Aufgabe darin, den für die aktuelle
25 Lichtsituation geeigneten 8 Bit Bereich innerhalb der 12 Bit zu bestimmen (z.B. die unteren 8 Bit bei Nacht und / oder Fahrt in einem Tunnel, die oberen 8 Bit bei hellem Sonnenschein, oder Fahrt in die untergehende Sonne). Bei dieser Methode bleibt die Belichtungszeit unverändert. Die Wahl des
30 Bereichs geschieht ebenfalls durch die Auswertung der Intensitätsverteilung in ganzen Bild und/oder Bildteilen unter CPU-Last.

Besonders vorteilhaft lässt sich jedoch die Belichtungsregelung unter Zurhilfenahme einer GPU, insbesondere mittels deren Pixel-Shadern, realisieren. Die GPU ist üblicherweise für die Verarbeitung von Farbbildern ausgelegt. Für die Farbbildverarbeitung wird eine RGBA-Format verwendet. Das bedeutet es stehen jeweils 8 Bit für Rot (R), Grün (G), Blau (B) und einen Alpha-Kanal (A) zur Verfügung (Der Grafikkarten Marktführer NVIDIA erwägt die 8 Bit pro Farbkanal bis zu 32 Bit pro Farbkanal zu erweitern). Wird in diesem Format ein Grauwertbild verarbeitet, so wird in jedem der drei Kanäle R, G, und B der selbe, den Grauwert repräsentierende Skalar abgelegt. Im Prinzip wird also das 8 Bit Grauwertbild dreifach in dem zur Verfügung stehenden Bildspeicher abgelegt. Für die fehlenden 4 Bit des Kamerabildes können nun die 8 Bit des Alpha-Kanals verwendet werden. Durch geeignete Implementierung werden aus den 8 Bit Grauwert im RGB und den 4 Bit Grauwert im Alpha-Kanal wieder die ganzen 12 Bit Grauwert der Kamera in der GPU hergestellt. Da standardmäßig am Display-Ausgang einer GPU die zur Darstellung zu bringenden RGB-Bilddaten als 8-Bit Information zur Verfügung gestellt werden, kann durch die entsprechende Auswahl der Zwischengespeicherten der 8 Bit RGB-Grauwertinformation der Kamera und der 4 Bit Alpha-Grauwertinformation der Kamera der gewünschte 8 Bit-Bereich gewählt werden, der angezeigt werden soll. Dies ist gleichbedeutend mit der Belichtungssteuerung der Kamera. Die Transformation der 12 Bit Kamera-Werte in das RGBA-Format geschieht hierbei in der Regel durch eine der GPU beigegebene CPU. Die CPU-Last ist bei Anwendung dieses vorteilhaften Verfahrens zur Kamera-Belichtungsregelung jedoch wesentlich geringer, als die bei der Belichtungssteuerung durch Intensitätsauswertung anfallende. Hinzukommt, dass der PCI- bzw. AGP-Bus für den DMA auf 32 Bit Datenübertragung optimiert ist. Beim RGBA-Format fallen genau diese 32 Bit an, so dass die dreifache Bildübertragung (R=G=B) gerade keinen Nachteil,

sondern einen Vorteil bietet. Darüber hinaus, können die Daten, die nicht angezeigt werden, mit 12 Bit in der GPU bearbeitet werden. Diese 12 Bit können nach der Bearbeitung in der GPU mit dem gleichen Verfahren der CPU zu deren Weiter-

5 verarbeitung zur Verfügung gestellt werden.

Patentansprüche

1. Vorrichtung zur Bildverarbeitung, bestehend aus
einer Einheit zur Verfügungsstellung von Bilddaten,
5 einer Rechneinheit zur Erkennung von Objekten in den Bilddaten und
einer Einheit zur Weiterverarbeitung der Erkennungsergebnisse,
dadurch gekennzeichnet,
10 dass die Rechneinheit zur Erkennung der Objekte im wesentlichen aus einem Graphik-Prozessor (GPU bzw. VPU), insbesondere mit einer im zugeordneten Hardware-Peripherie, besteht.
2. Vorrichtung nach Anspruch 1,
15 dadurch gekennzeichnet,
dass die Einheit zur Verfügungsstellung der Bilddaten im wesentlichen aus einem Kamerasystem und/oder einem Bildspeicher besteht.
- 20 3. Vorrichtung nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet,
dass die Einheit zur Weiterverarbeitung der Erkennungsergebnisse im wesentlichen aus einer Bildanzeige und/oder einer
Einheit besteht, welche auf Grund der in den Bilddaten er-
25 kannten Objekte auf andere Systeme einwirkt.
4. Vorrichtung nach einem der vorhergehenden Ansprüche,
dadurch gekennzeichnet

dass zur Unterstützung bei den Bildverarbeitungsaufgaben parallel zu dem Graphikprozessor eine frei programmierbare CPU angeordnet ist.

- 5 5. Verfahren zur Bildverarbeitung,
bei welchem Bilddaten zur Verarbeitung zur Verfügung gestellt
werden,
worauf in diesen Bilddaten mittels einer Rechneinheit Ob-
jekte erkannt werden,
10 und die Erkennungsergebnisse an eine Einheit zur Weiterver-
arbeitung weitergeleitet werden,
d a d u r c h g e k e n n z e i c h n e t ,
dass die Algorithmen zur Erkennung der Objekte in den Bildda-
ten im wesentlichen auf einem Graphik-Prozessor (GPU bzw.
15 VPU), insbesondere unter Zurhilfenahme der diesem zugeordne-
ten Hardware-Peripherie, ablaufen.

6. Verfahren nach Anspruch 5,
d a d u r c h g e k e n n z e i c h n e t ,
20 dass die Bilddaten durch ein Kamerasystem und/oder einen
Bildspeicher zur Verfügung gestellt werden.

7. Verfahren nach einem der vorhergehenden Ansprüche,
d a d u r c h g e k e n n z e i c h n e t ,
25 dass im Rahmen der Weiterverarbeitung der Erkennungsergebnis-
se diese auf einer Anzeige dargestellt werden und/oder einem
System zur Verfügung gestellt werden, welches auf Grund der
in den Bilddaten erkannten Objekte auf andere Systeme ein-
wirkt.

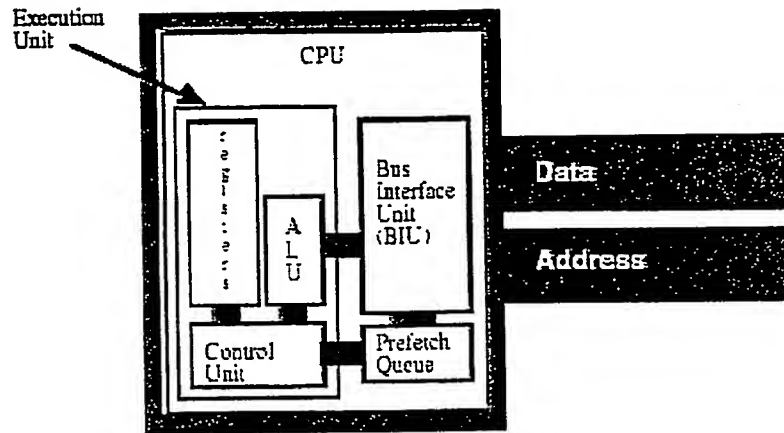
30

8. Verfahren nach einem der vorhergehend Ansprüche,
d a d u r c h g e k e n n z e i c h n e t ,

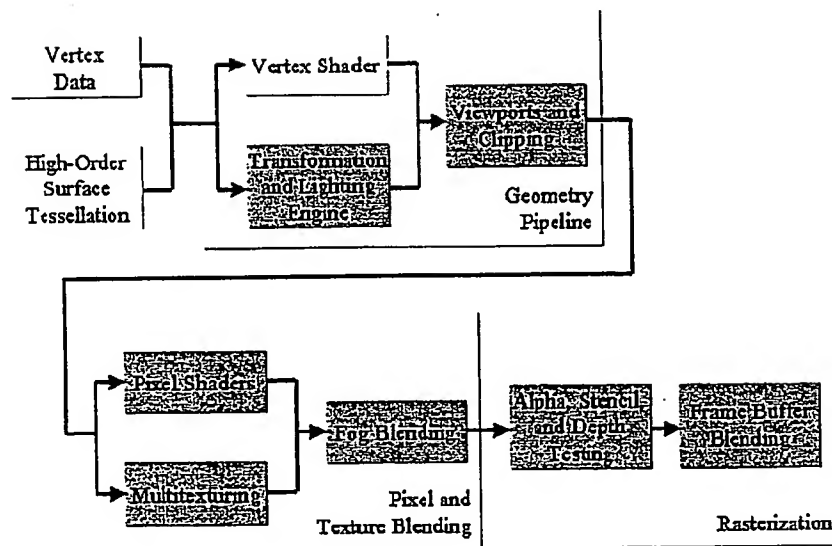
dass die GPU mit einer CPU in Verbindung steht, welche in steuernder oder regelnder Weise auf den Verarbeitungsablauf in der GPU einwirkt.

- 5 9. Verfahren nach Anspruch 8,
 d a d u r c h g e k e n n z e i c h n e t ,
 dass die Bilddaten via DMA (Direkt Memory Access) von der Kamera oder einem Bildspeicher direkt in die GPU und nach der
 Aufbereitung durch die GPU durch DMA direkt in den Hauptspeicher
10 cher der CPU zur Weiterverarbeitung geladen werden, und dieses Vorgehen gegebenenfalls auf umgekehrt und/oder iterativ erfolgt.

1/1



Figur 1
(01920) XNA 18 30A9 21HT



Figur 2

BEST AVAILABLE COPY

THIS PAGE BLANK (USPTO)